Ruoying Hao

February 11, 2017

I tried to install Tensorflow on a PC laptop. But immediately realized that it needs Anaconda platform to operate. I’ll try to install the platform on Sunday. I may also try to see if the platform is installed on school Linux machine.

I have been successful in writing the code to massively download the 10-Qs for S&P 100. So far I have the Python code on my hand that, if given a ticker list text file, it can find the latest 20 entries for 10-Q interactive data page, and download the excel file onto a local drive. The excel files, originally named uniformly Financial\_Report.xlsx, would be renamed following the format TICKER\_DATE\_10Q.xlsx. The date is the filing date, instead of the date that indicate the end of the quarter. The filing date is extracted from the pages that have the 10-Q list by using Beautifulsoup package.

The filing date is important because a company generally make a public disclosure of their quarterly financial condition several days BEFORE the filing date.

I’ve tested the code in various way but have not started download the 1600 excel files. I would do it on a school Linux machine. Hopefully it is faster than doing it on my laptop.

I have created Dropbox and Bitbucket account. As long as we are connected, we can start sharing files.